



**QUEEN'S
UNIVERSITY
BELFAST**

Joint Load Balancing and Interference Management for Small-Cell Heterogeneous Networks with Limited Backhaul Capacity

Tam, H. H. M., Tuan, H. D., Ngo, D. T., Duong, T. Q., & Poor, H. V. (2016). Joint Load Balancing and Interference Management for Small-Cell Heterogeneous Networks with Limited Backhaul Capacity. *IEEE Transactions on Wireless Communications*, 16(2), 872. <https://doi.org/10.1109/TWC.2016.2633262>

Published in:

IEEE Transactions on Wireless Communications

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Joint Load Balancing and Interference Management for Small-Cell Heterogeneous Networks with Limited Backhaul Capacity

Ho H. M. Tam, Hoang D. Tuan, Duy T. Ngo, Trung Q. Duong and H. Vincent Poor

Abstract—In this paper, new strategies are devised for joint load balancing and interference management in the downlink of a heterogeneous network, where small cells are densely deployed within the coverage area of a traditional macrocell. Unlike existing work, the limited backhaul capacity at each base station (BS) is taken into account. Here users (UEs) cannot be offloaded to any arbitrary BS, but only to ones with sufficient backhaul capacity remaining. Jointly designed with traffic offload, transmit power allocation mitigates the intercell interference to further support the quality of service of each UE. The objective here is either (i) to maximize the network sum rate subject to minimum throughput requirements at individual UEs, or (ii) to maximize the minimum UE throughput. Both formulated problems belong to the difficult class of mixed-integer nonconvex optimization. The inherently binary BS-UE association variables are strongly coupled with the transmit power variables, making the problems even more challenging to solve. New iterative algorithms are developed based on an exact penalty method combined with successive convex programming, where we deal with the binary BS-UE association problem and the nonconvex power allocation problem one at a time. At each iteration of the proposed algorithms, only two simple convex problems need to be solved in the same time scale. It is proven that the algorithms improve the objective functions in each iteration and converge eventually. Numerical results demonstrate the efficiency of the proposed algorithms in both traffic offloading and interference mitigation.

Index Terms—Combinatorial optimization, heterogeneous network, limited backhaul, load balancing, power allocation, successive convex programming

Manuscript received May 21, 2016; revised August 22, 2016; accepted November 18, 2016. The associate editor coordinating the review of this paper and approving it for publication was S. Mukherjee.

H. H. M. Tam and H. D. Tuan are with the Faculty of Engineering and Information Technology, University of Technology Sydney, Broadway, NSW 2007, Australia (email: HuuMinhTam.Ho@student.uts.edu.au, Tuan.Hoang@uts.edu.au).

D. T. Ngo is with the School of Electrical Engineering and Computing, The University of Newcastle, Callaghan, NSW 2308, Australia (email: duy.ngo@newcastle.edu.au).

T. Q. Duong is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT7 1NN, United Kingdom (e-mail: trung.q.duong@qub.ac.uk).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

The work of H. H. M. Tam and H. D. Tuan was supported in part by the Australian Research Council's Discovery Projects under Project DP130104617. The work of T. Q. Duong was supported in part by the U.K. Royal Academy of Engineering Research Fellowship under Grant RF1415\14\22 and by the Newton Institutional Link under Grant ID 172719890.

I. INTRODUCTION

Cell densification is currently the best hope to meet the unprecedented data increase (the $1000\times$ data challenge) in the fifth-generation (5G) wireless networks [1]–[3]. By densely deploying cells of different types and sizes (e.g., macro, micro, pico, femto), the resulting heterogeneous networks (HetNets) can offer a substantial growth in area spectral efficiency and full network coverage in regions traditionally difficult to penetrate. Another key benefit of HetNet is data offloading, where traffic otherwise transported via the traditional macrocell is directed to the newly deployed small cells.

Traditionally, a user (UE) is associated with the base station (BS) that offers the maximum signal-to-interference-plus noise ratio (SINR), i.e., the max-SINR rule (see, e.g., [4]). As a result, a 'hotspot' BS with advantageous link conditions and/or high transmit power would potentially be inundated with too many UEs while other BSs only serve a few UEs. Range expansion is a heuristic method that may help balance the traffic load among different BSs, in which the SINR is regulated through a positive bias level [5], [6]. Still, it is challenging to determine optimal bias levels for multiple cells. Other related BS-UE association rules proposed in the literature are based on maximizing the estimated throughput [7] and sum logarithmic throughput [8]–[10]. Using Lagrangian duality decomposition [11], the association rule in [12] aims at maximizing the network sum-rate while satisfying the Quality-of-Service (QoS) constraints. A heuristic adjustment is then proposed to keep the total number of time slots demanded by the UEs below that available at the BSs. In [13], [14], a binary relaxation method is proposed to find the optimal association rule for the sum-rate and minimum-rate maximization objectives. However, the proposed method is limited to a two-cell network. An extensive overview of the state-of-the-art in user association for 5G networks can be found in [15].

A common assumption in the above existing work is the availability of ideal backhaul links with unlimited capacity. This assumption is not true for HetNets. Here, the low-power BSs of small cells (e.g., pico and femto) connect to the core network via low capacity backhaul (e.g., DSL links) for economic benefits [16]. If these small-cell BSs must serve too many UEs, their non-ideal backhaul becomes the bottleneck in transporting the required amount of data traffic to the UEs, resulting in a potentially unacceptable level of delay/jitter at the UEs. The study of [17] proposes an optimal BS-UE association rule that maximizes the logarithmic utility function

while guaranteeing a target delay. Note that one can guarantee certain levels of delay if the demanded throughput at the BSs is kept below their respective backhaul capacity [18]. Instead of maximizing the network sum-rate or the minimum UE's throughput, [19] devises a backhaul-aware BS-UE association rule that is based on biasing, cell size and user distribution.

Interference is a major issue in dense small-cell HetNets, wherein numerous cell boundaries with poorly defined patterns are created. Compared with traditional cellular networks, the effects of the intercell interference are much more acute and unpredictable, especially at the cell edges [20]. Power control is an effective way to manage the interference, assuming that the BS-UE links have already been established. For a given BS-UE association, [21] devises an optimal BS transmit power policy for sum-rate maximization with backhaul capacity constraints by solving the Karush-Kuhn-Tucker (KKT) conditions. However, due to the nonconvexity of the considered problem, a KKT point may not even be locally optimal or feasible. In addition, the once presumably optimal BS-UE associations will no longer be optimal when the new transmit power values are used as a result of power control. It is therefore essential to design jointly optimal strategies for both traffic offload and interference management.

For CDMA-based networks, joint optimization of BS-UE association and interference management is considered in [22] for network sum-rate maximization and in [23] for minimum UE's SINR maximization. It is not straightforward to apply the results of [22] and [23] to networks in which a BS uses orthogonal channels to serve its UEs to eliminate intracell interference. Different from CDMA, each UE here is only assigned with a fraction of the time/frequency slots depending on the current load at its serving BS. Assuming zero intracell interference, [24] proposes an iterative procedure for joint BS-UE association and interference management that guarantees a maximum delay not be exceeded. Yet, the convergence of the proposed heuristic method is not proven. Using game theory, [25] finds the Nash equilibrium for such joint optimization problem, albeit without considering QoS constraints. It is commonly known that a Nash equilibrium may not be efficient as it could be far away from the actual optimal solution. Notably, the practical issue of imperfect backhaul links is not considered in [24] and [25], presumably due to the nonconvexity of the backhaul capacity constraints even when the BS-UE association is fixed.

In this paper, we formulate new problems for joint traffic offload and interference management in the downlink of a HetNet. Aiming to maximize the network throughput and the minimum UE rate, our formulations accommodate both backhaul capacity constraints and UE QoS requirements. The considered problems belong to the difficult class of mixed integer nonconvex optimization. The binary BS-UE association variables are strongly coupled with the transmit power variables, making the problems even more challenging to solve.

We then develop new iterative algorithms based on alternating descent [26] and successive convex programming for the formulated problems. Alternating descent allows us to decouple the original problem into two subproblems and deal

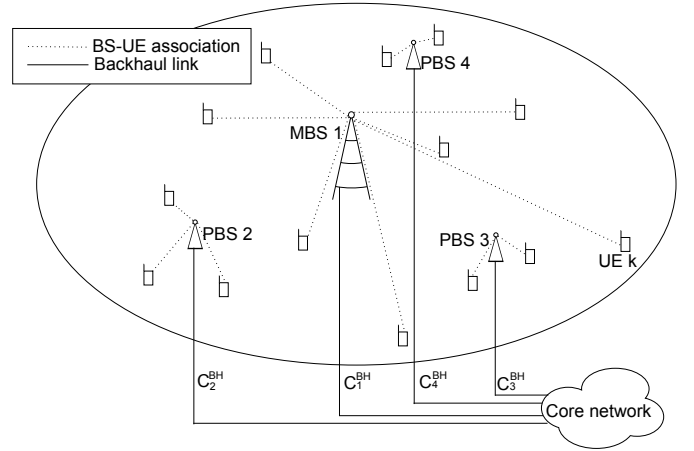


Fig. 1. A small-cell HetNet with limited-capacity backhaul links. 'MBS', 'PBS' and 'PoP' refer to macro BS, pico BS and Point of Presence, respectively.

with them one at a time. Even so, each resulting subproblem is still challenging. For a fixed power allocation, the BS-UE association subproblem is combinatorial. And for a fixed BS-UE association, the power allocation subproblem is highly nonconvex. We propose to deal with the binary nature of BS-UE association by relaxation combined with a penalty method. We then employ successive convex programming to solve two subproblems in the same time scale. We prove that our proposed alternating descent algorithms converge, where only two simple convex problems are to be solved in each iteration. Simulation results show that the proposed algorithms enhance the network throughput through better load balancing and interference management.

The rest of this paper is organized as follows: Sec. II formulates the problems of joint load balancing and interference management. Sec. III proposes an alternating descent algorithm to solve the sum-rate maximization problem. Sec. IV extends the devised solution to the case of minimum UE rate maximization. Sec. V presents numerical results to demonstrate the performance of our proposed algorithms. Finally, Sec. VI concludes the paper.

Notation. In this paper, boldfaced symbols are used for optimization variables whereas non-boldfaced symbols are for deterministic terms, regardless of whether they are matrix, vector or scalar. The dimensions of these symbols are interpreted from context, and they will be explicitly specified should there be any potential ambiguity.

II. SYSTEM MODEL AND PROBLEM FORMULATIONS

Consider the downlink of a K -cell HetNet in which one macrocell is overlaid with $K - 1$ small cells, as depicted in Fig. 1. To best exploit the limited radio spectrum, universal frequency reuse is adopted [8], [10]. Without loss of generality, we assume that macro base station (MBS) serving the macro-cell is indexed as BS 1, and the BS serving small cell k (e.g., a micro/pico/femto BS) is indexed as BS $k \in \{2, 3, \dots, K\}$. Transmitting at a much lower power than an MBS, the small-cell BSs are deployed densely in order to extend network

coverage, increase throughput and offload data traffic from the MBS.

We assume that each BS $k \in \{1, 2, \dots, K\}$ connects to a backhaul link with limited capacity. The backhaul link of each BS carries the downlink traffic for its serviced UEs from the core network to that BS via a central access point called ‘Point of Presence’ (PoP) [17], [27], [28]. The PoP is connected to the core network via an optical fiber link whose capacity is much higher than the total capacity of all links from the PoP to all BSs. Therefore, the effect of core network-PoP link capacity is neglected in our model [28]. For simplicity, we also neglect the traffic coming from the control plane [18].

In the considered HetNet, there are N UEs looking for the serving BSs. Similar to the BSs, the UEs are each equipped with one antenna. A snapshot model is adopted where the channels remain unchanged during the optimization process. This channel assumption is well-justified for networks with a low degree of mobility and/or very high throughput. A central processing unit is employed to collect all the channel state information and perform the underlying network optimization.

In this paper, a UE is allowed to associate with at most one BS, but a BS can serve multiple UEs. Assume that BS $k \in \{1, 2, \dots, K\}$ has a full buffer and transmits with power \mathbf{p}_k . First, consider that only one UE $n \in \{1, 2, \dots, N\}$ is connected to BS k . The achieved data rate in nats/s/Hz of UE n is expressed as:

$$r_{nk}(\mathbf{p}) \triangleq \ln \left(1 + \frac{g_{nk}\mathbf{p}_k}{\sum_{j \neq k, j=1}^K g_{nj}\mathbf{p}_j + \sigma^2} \right), \quad (1)$$

where $\mathbf{p} \triangleq (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K)^T$, g_{nk} is the channel gain from BS k to UE n , and σ^2 is the power of background additive white Gaussian noise. As seen from (1), UE n is subjected to the intercell interference from other BS $j \neq k$.

Next, consider the general case of multiple UEs connecting to a BS. The BS will then divide the total available time into a number of time slots and allocates them to its serviced UEs in a round-robin fashion [8], [10]. As such, each connected UE will receive an equal amount of transmission time while there is no intracell interference. Denote $\mathbf{x}_{nk} \in \{0, 1\}$ as the BS-UE association variable, i.e., $\mathbf{x}_{nk} = 1$ if UE n is associated with BS k and $\mathbf{x}_{nk} = 0$ otherwise. Define $\mathbf{x}_k \triangleq [\mathbf{x}_{1k}, \dots, \mathbf{x}_{Nk}]^T$ and $\mathbf{x} \triangleq [\mathbf{x}_1^T, \dots, \mathbf{x}_K^T]^T$. If BS k serves a total of $\langle \mathbf{x}_k \rangle \triangleq \sum_{n=1}^N \mathbf{x}_{nk}$ UEs, then each of these UEs will be allocated $1/\langle \mathbf{x}_k \rangle$ of the total available time. Effectively, data rate perceived by a connected UE n is $r_{nk}(\mathbf{p})/\langle \mathbf{x}_k \rangle$, which will be further reduced as more and more UEs are associated to BS k . To reflect the fact that this rate is only possible if UE n actually connects to BS k , we define the effective data rate given to UE n by BS k as:

$$r_{nk}^{\text{eff}}(\mathbf{x}_k, \mathbf{p}) \triangleq \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle}. \quad (2)$$

It follows that the sum effective data rate of cell k is

$$r_k^{\text{eff}}(\mathbf{x}, \mathbf{p}) \triangleq \sum_{n=1}^N r_{nk}^{\text{eff}}(\mathbf{p}) = \sum_{n=1}^N \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle}, \quad (3)$$

which is required not to exceed the limited backhaul capacity available to BS k [29], [30]. The total network throughput across all K cells is then simply

$$\sum_{k=1}^K r_k^{\text{eff}}(\mathbf{x}, \mathbf{p}) = \sum_{k=1}^K \sum_{n=1}^N \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle}. \quad (4)$$

This work aims to enhance the network throughput by joint optimization of BS-UE association and transmit power allocation. Importantly, our design takes into account both the QoS requirement of each UE and the limited backhaul capacity at each BS. If too many UEs connect a particular BS (e.g., due to favorable channel conditions), then (i) the perceived rate of each UE will decrease and potentially not satisfy the minimum QoS requirement, and (ii) the sum effective rate of the corresponding cell may increase and potentially exceed the backhaul capacity. With a proper BS-UE association, the traffic load will be more balanced among different BSs and the network crowding issue can be alleviated. With adaptive power allocation, the intercell interference can be effectively managed to further improve the throughput. Here, we will consider the following joint design problem for sum-rate maximization.

$$\max_{\mathbf{p}, \mathbf{x}_{nk} \in \{0, 1\}} \sum_{k=1}^K \sum_{n=1}^N \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle} \quad (5a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \mathbf{x}_{nk} = 1, \quad n = 1, \dots, N \quad (5b)$$

$$\langle \mathbf{x}_k \rangle \geq 1, \quad k = 1, \dots, K \quad (5c)$$

$$\sum_{k=1}^K \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle} \geq R_n^{\text{QoS}}, \quad n = 1, \dots, N \quad (5d)$$

$$\sum_{n=1}^N \mathbf{x}_{nk}r_{nk}(\mathbf{p}) \leq C_k^{\text{BH}}\langle \mathbf{x}_k \rangle, \quad k = 1, \dots, K \quad (5e)$$

$$0 \leq \mathbf{p}_k \leq P_k^{\text{max}}, \quad k = 1, \dots, K. \quad (5f)$$

Constraint (5b) ensures each UE be connected with one BS only. Constraint (5c) requires each BS serves at least one UE [8]–[10], [17]. In (5d) and (5e), $R_n^{\text{QoS}} \geq 0$ and $C_k^{\text{BH}} \geq 0$ specify the minimum throughput requirement for UE n and the backhaul capacity for cell k , respectively. Finally, (5f) caps the maximum transmit power of each BS k .

We also consider the following problem of maximizing the minimum effective rate among all UEs:

$$\max_{\mathbf{p}, \mathbf{x}_{nk} \in \{0, 1\}} \min_{n=1, \dots, N} \left\{ \sum_{k=1}^K \frac{\mathbf{x}_{nk}r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k \rangle} \right\} \quad (6)$$

s.t. (5b), (5c), (5e), (5f).

In problem (6), our aim is to support the most vulnerable UEs, e.g., those at the cell edges.

Both problems (5) and (6) belong to the difficult class of mixed-integer programming. The strong coupling between the binary variables \mathbf{x} and the continuous variables \mathbf{p} make the problems even more challenging. State-of-the-arts in existing literature typically apply the alternating optimization framework [10], [22], [23], [31], [32]. In this ‘divide-and-conquer’

approach, instead of dealing with \mathbf{x} and \mathbf{p} simultaneously, one decouples the original problems into subproblems of lower dimensions and resolve one subproblem at a time. Still for our problems (5) and (6) at hand, the BS-UE association problem for optimization in binary \mathbf{x} scales exponentially with the number of BSs and UEs. It is not practical to try all the possible BS-UE combinations, even for networks of small-to-medium size. Moreover, for a given BS-UE combination, the power allocation for optimization in \mathbf{p} remains highly nonconvex. Specifically, problem (5) has a nonconvex objective subject to nonconvex QoS and backhaul constraints, whereas problem (6) has a nonsmooth nonconvex objective subject also to a nonconvex set.

In what follows, we will address both problems (5) and (6) by a novel alternating descent method, which aims at improving the iterative solutions. It is noteworthy that the proposed joint user association and power control algorithms, although designed for single-antenna networks, can serve as a fundamental building block for subsequent development of joint user association and beamforming/precoding in multiple-antenna networks [10].

III. PROPOSED ALTERNATING DESCENT ALGORITHM FOR SUM-RATE MAXIMIZATION

A. BS-UE Association for Fixed Transmit Power

Given a fixed $\mathbf{p} := p$, we aim to solve problem (5) in variable \mathbf{x} . References [13], [14] considered the simplest case with $K = 2$ cells, under which the objective function (5a)

$$\begin{aligned} & \frac{\sum_{n=1}^N \mathbf{x}_{n1} r_{n1}(p)}{\langle \mathbf{x}_1 \rangle} + \frac{\sum_{n=1}^N \mathbf{x}_{n2} r_{n2}(p)}{\langle \mathbf{x}_2 \rangle} \\ &= \frac{\sum_{n=1}^N \mathbf{x}_{n1} r_{n1}(p) \langle \mathbf{x}_2 \rangle + \sum_{n=1}^N \mathbf{x}_{n2} r_{n2}(p) \langle \mathbf{x}_1 \rangle}{\langle \mathbf{x}_1 \rangle \langle \mathbf{x}_2 \rangle} \end{aligned}$$

is a fraction of linear functions in the new rank-one constrained matrix variable $\mathbf{X} = \mathbf{x}\mathbf{x}^T$ for $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$. By dropping the constraint $\text{rank}(\mathbf{X}) = 1$ and relaxing binary constraints on its entries to real numbers belonging to the interval $[0, 1]$, a bisection search is used in finding the optimal solution of the resultant program. It should be emphasized again that such relaxation only works if there are two cells in the network. This is because for $K > 2$ the objective function (5a) becomes a fraction of nonlinear functions in $\mathbf{X} = \mathbf{x}\mathbf{x}^T$.

Our objective here is to devise a solution that works for a general network with an arbitrary number of cells. To begin with, notice that all the constraints (5b), (5c) and (5e) are linear in \mathbf{x} , but not (5d). The following proposition allows us to equivalently recast the nonconvex constraint (5d) as a system of linear constraints on \mathbf{x} .

Proposition 1: Under the constraint (5b), the constraint (5d) is equivalent to

$$(M - (M - 1)\mathbf{x}_{nk}) r_{nk}(p) \geq R_n^{\text{QoS}} \langle \mathbf{x}_k \rangle, \quad (7)$$

$$n = 1, \dots, N, \quad k = 1, \dots, K$$

for a sufficiently large number M .

Proof: Denote by $(5d)_n$ the constraint (5d) for n . It is sufficient to show that each $(5d)_n$ is equivalent to the following

K constraints:

$$(M - (M - 1)\mathbf{x}_{nk}) r_{nk}(p) \geq R_n^{\text{QoS}} \langle \mathbf{x}_k \rangle, \quad k = 1, \dots, K. \quad (8)$$

Under the constraint (5b), for each n there is k_n such that $\mathbf{x}_{nk_n} = 1$ and $\mathbf{x}_{nk} = 0, \forall k \in \{1, \dots, K\} \setminus \{k_n\}$. Therefore, (8) merely means that

$$r_{nk_n}(p) \geq R_n^{\text{QoS}} \langle \mathbf{x}_{k_n} \rangle, \quad (9)$$

and

$$Mr_{nk}(p) \geq R_n^{\text{QoS}} \langle \mathbf{x}_k \rangle, \quad k \neq k_n. \quad (10)$$

Note that (9) is $(5d)_n$, showing the implication $(8) \Rightarrow (5d)_n$.

On the other hand, (10) holds true for $0 < M < +\infty$ because its right hand side is obviously bounded while the factor $r_{nk}(p)$ on the left hand side is strictly positive. The inverse implication $(5d)_n \Rightarrow (8)$ thus follows, yielding the equivalence between (8) and $(5d)_n$. ■

Next, we deal with the binary nature of \mathbf{x} . For $\mathbf{x}_{\hat{n}k} \in \{0, 1\}$, one has $\mathbf{x}_{\hat{n}k} = \mathbf{x}_{\hat{n}k}^2$ and thus $\langle \mathbf{x}_k \rangle = \langle \mathbf{x}_k^2 \rangle \triangleq \sum_{\hat{n}=1}^N \mathbf{x}_{\hat{n}k}^2$. The objective function (5a) is then expressed as

$$\sum_{k=1}^K \sum_{n=1}^N \frac{\mathbf{x}_{nk}^2 r_{nk}(p)}{\langle \mathbf{x}_k^2 \rangle}. \quad (11)$$

On one hand, it is straightforward to see that $\mathbf{x}_{nk} \in \{0, 1\}$ is equivalent to $\mathbf{x}_{nk}^2 = \mathbf{x}_{nk}$, $\mathbf{x}_{nk} \in [0, 1]$. On the other hand, it holds true that $\mathbf{x}_{nk}^2 \leq \mathbf{x}_{nk}$ for $\mathbf{x}_{nk} \in [0, 1]$. Following [33], we relax binary \mathbf{x}_{nk} to $\mathbf{x}_{nk} \in [0, 1]$ and introduce a penalty term in the objective function (11) to enforce $\mathbf{x}_{nk}^2 = \mathbf{x}_{nk}$, thus making \mathbf{x}_{nk} binary. This leads to the following problem:

$$\begin{aligned} & \max_{\mathbf{x}_k \in [0, 1]^N} \mathcal{P}_1(\mathbf{x}, p) \triangleq \sum_{k=1}^K \sum_{n=1}^N \frac{\mathbf{x}_{nk}^2 r_{nk}(p)}{\langle \mathbf{x}_k^2 \rangle} \\ & \quad + \lambda \sum_{k=1}^K \sum_{n=1}^N (\mathbf{x}_{nk}^2 - \mathbf{x}_{nk}) \\ & \text{s.t.} \quad (5b) - (5e), \end{aligned} \quad (12)$$

where $\lambda \geq 0$ is a constant penalty factor. Parameter λ signifies the relative importance of recovering binary values for \mathbf{x} over throughput maximization. In (12), the term $\sum_{k=1}^K \sum_{n=1}^N (\mathbf{x}_{nk} - \mathbf{x}_{nk}^2)$ is always nonnegative and can therefore be used to measure the degree of satisfaction of the binary constraints $\mathbf{x}_{nk} \in \{0, 1\}$, $\forall n, k$. Without squaring such a term, the above penalization is exact, meaning that the constraints $\mathbf{x} \in \{0, 1\}$, $\forall n, k$ can be satisfied by a maximizer of (12) with a finite value of λ (see, e.g., [34, Ch. 16]). This nice property makes such exact penalization attractive.

With Proposition 1, problem (12) becomes

$$\begin{aligned} & \max_{\mathbf{x}_k \in [0, 1]^N} \mathcal{P}_1(\mathbf{x}, p) \\ & \text{s.t.} \quad (5b), (5c), (5e), (7). \end{aligned} \quad (13)$$

With an appropriate choice of λ , problems (13) and (5) are equivalent in the sense that they share the same optimal solution [33, Sec. II]. Since problem (13) is still nonconvex, we now employ successive convex programming to solve it.

Proposition 2: For a given point $(x^{(\kappa)}, p)$, the following convex problem is a global lower bound maximization for (13):

$$\begin{aligned} \max_{\mathbf{x}_k \in [0,1]^N} \quad & \tilde{\mathcal{P}}_1^{(\kappa)}(\mathbf{x}, p) \triangleq \sum_{k=1}^K \sum_{n=1}^N \alpha_{nk}^{(\kappa)}(\mathbf{x}, p) \\ & + \lambda \sum_{k=1}^K \sum_{n=1}^N \gamma_{nk}^{(\kappa)}(\mathbf{x}) \\ \text{s.t.} \quad & (5b), (5c), (5e), (7), \end{aligned} \quad (14)$$

where we define

$$\begin{aligned} \alpha_{nk}^{(\kappa)}(\mathbf{x}, p) \triangleq & \frac{(x_{nk}^{(\kappa)})^2 r_{nk}(p)}{\langle (x_k^{(\kappa)})^2 \rangle} + \frac{2x_{nk}^{(\kappa)} (\mathbf{x}_{nk} - x_{nk}^{(\kappa)}) r_{nk}(p)}{\langle (x_k^{(\kappa)})^2 \rangle} \\ & - \frac{(x_{nk}^{(\kappa)})^2 r_{nk}(p)}{\left(\langle (x_k^{(\kappa)})^2 \rangle \right)^2} \left(\langle \mathbf{x}_k^2 \rangle - \langle (x_k^{(\kappa)})^2 \rangle \right), \quad (15a) \end{aligned}$$

$$\gamma_{nk}^{(\kappa)}(\mathbf{x}) \triangleq (x_{nk}^{(\kappa)})^2 - x_{nk}^{(\kappa)} + (2x_{nk}^{(\kappa)} - 1) (\mathbf{x}_{nk} - x_{nk}^{(\kappa)}). \quad (15b)$$

Proof: See Appendix A, where we show that the objective $\tilde{\mathcal{P}}_1^{(\kappa)}(\cdot, p)$ in (14) is a global lower bound of the objective $\mathcal{P}_1(\mathbf{x}, p)$ in (13), i.e.

$$\mathcal{P}_1(\mathbf{x}, p) \geq \tilde{\mathcal{P}}_1^{(\kappa)}(\mathbf{x}, p), \quad \forall \mathbf{x} \text{ and } \mathcal{P}_1(x^{(\kappa)}, p) = \tilde{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa)}, p). \quad (16)$$

The nonconvex problem (13) can then be addressed by instead solving its global lower bound maximization (14) in a sequential manner as follows: After initializing from a feasible point $x^{(0)}$ of problem (13), we iteratively solve problem (14) to generate a sequence $\{x^{(\kappa)}\}$, $\kappa = 1, 2, \dots$ of feasible and improved points toward the optimal solution of (13). More specifically, at iteration κ we use $x^{(\kappa-1)}$ as a feasible point to solve (14) and obtain $x^{(\kappa)}$.

Theorem 1: Initialized from a feasible point $x^{(0)}$, the sequence $\{x^{(\kappa)}\}$ obtained by iteratively solving (14) is a sequence of improved points of (13), which converges to a KKT point.

Proof: Note that $x^{(\kappa)}$ and $x^{(\kappa+1)}$ are a feasible point and the optimal solution of (14), respectively. By using (16),

$$\begin{aligned} \mathcal{P}_1(x^{(\kappa+1)}, p) & \geq \tilde{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa+1)}, p) \geq \tilde{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa)}, p) \\ & = \mathcal{P}_1(x^{(\kappa)}, p), \end{aligned} \quad (17)$$

i.e. $x^{(\kappa+1)}$ is a better point of (13) than $x^{(\kappa)}$. Since the sequence $\{x^{(\kappa)}\}$ is bounded, by Cauchy's theorem there is a convergent subsequence $\{x^{(\kappa_\nu)}\}$ with a limit point \bar{x} , i.e.

$$\lim_{\nu \rightarrow +\infty} [\mathcal{P}_1(x^{(\kappa_\nu)}, p) - \mathcal{P}_1(\bar{x}, p)] = 0.$$

For every κ there is ν such that $\kappa_\nu \leq \kappa \leq \kappa_{\nu+1}$ so

$$\begin{aligned} 0 & = \lim_{\nu \rightarrow +\infty} [\mathcal{P}_1(x^{(\kappa_\nu)}, p) - \mathcal{P}_1(\bar{x}, p)] \\ & \leq \lim_{\nu \rightarrow +\infty} [\mathcal{P}_1(x^{(\kappa)}, p) - \mathcal{P}_1(\bar{x}, p)] \\ & \leq \lim_{\nu \rightarrow +\infty} [\mathcal{P}_1(x^{(\kappa_{\nu+1})}, p) - \mathcal{P}_1(\bar{x}, p)] = 0, \end{aligned}$$

showing that $\lim_{\kappa \rightarrow +\infty} \mathcal{P}_1(x^{(\kappa)}, p) = \mathcal{P}_1(\bar{x}, p)$. Then, each accumulation point \bar{x} of the sequence $\{x^{(\kappa)}\}$ is a KKT-point according to [35, Theorem 1]. The proof of Theorem 1 is thus complete. ■

B. Power Allocation for Fixed BS-UE Association

Given $\mathbf{x} := x$, we proceed to solving problem (5) in the variable \mathbf{p} . Although the difference-of-convex iterations (DCI) approach of [36] can be applied in the absence of backhaul constraints (5e), the required log-determinant optimization is computationally expensive even for commercialized convex solvers. This drawback is particularly severe in HetNets which consist of a large number of densely deployed BSs and UEs. Reference [21] proposes solving the KKT conditions, followed by applying the gradient descent method to update Lagrangian multipliers in order to satisfy (5e). Nevertheless, a solution derived from the KKT conditions of a nonconvex problem may not be locally optimal or even feasible.

Our aim here is to devise an efficient and optimal power allocation solution. Using Proposition 1 and simple algebraic manipulations, (5d) is expressed as the following linear constraints:

$$\begin{aligned} g_{nk} \mathbf{p}_k & \geq \left[\exp \left(\frac{R_n^{\text{QoS}} \langle x_k \rangle}{M - (M-1)x_{nk}} \right) - 1 \right] \times \\ & \left(\sum_{j \neq k, j=1}^K g_{nj} \mathbf{p}_j + \sigma^2 \right), \quad n = 1, \dots, N, \quad k = 1, \dots, K. \end{aligned} \quad (18)$$

Problem (5) is then reduced to

$$\max_{\mathbf{p}} \mathcal{P}_1(x, \mathbf{p}) \triangleq \sum_{k=1}^K \sum_{n=1}^N \frac{(x_{nk})^2}{\langle x_k^2 \rangle} r_{nk}(\mathbf{p}) \quad (19a)$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{n=1}^N x_{nk} r_{nk}(\mathbf{p}) \leq C_k^{\text{BH}} \langle x_k \rangle, \quad k = 1, \dots, K \\ & (5f), (18). \end{aligned} \quad (19b)$$

Because (19a) and (19b) are still nonconvex in \mathbf{p} , we instead consider their convex bounds as given in the following result.

Proposition 3: The rate function $r_{nk}(\mathbf{p})$ in (1) admits

$$\begin{aligned} r_{nk}(\mathbf{p}) & \leq \theta_{nk}^{(\kappa)}(\mathbf{p}) \\ & \triangleq r_{nk}(p^{(\kappa)}) + \frac{1}{\sum_{j \neq k, j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2} \\ & \times \sum_{j \neq k, j=1}^K (g_{nj} p_j^{(\kappa)})^2 \left(\frac{1}{g_{nj} \mathbf{p}_j} - \frac{1}{g_{nj} p_j^{(\kappa)}} \right) \\ & + \frac{1}{\sum_{j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2} \sum_{j=1}^K g_{nj} (\mathbf{p}_j - p_j^{(\kappa)}) \end{aligned} \quad (20)$$

as its upper bound, and

$$\begin{aligned}
 r_{nk}(\mathbf{p}) &\geq \beta_{nk}^{(\kappa)}(\mathbf{p}) \\
 &\triangleq r_{nk}(p^{(\kappa)}) - \frac{1}{\sum_{j=1}^K g_{nj}p_j^{(\kappa)} + \sigma^2} \\
 &\quad \times \sum_{j=1}^K (g_{nj}p_j^{(\kappa)})^2 \left(\frac{1}{g_{nj}p_j} - \frac{1}{g_{nj}p_j^{(\kappa)}} \right) \\
 &\quad - \frac{1}{\sum_{j \neq k, j=1}^K g_{nj}p_j^{(\kappa)} + \sigma^2} \sum_{j \neq k, j=1}^K g_{nj}(\mathbf{p}_j - p_j^{(\kappa)})
 \end{aligned} \tag{21}$$

as its lower bound.

Proof: See Appendix B. ■

With the bounds in (20) and (21), we now address the nonconvex problem (19) by successive convex programming. Specifically, after initializing a feasible point $p^{(0)}$ of problem (19), we iteratively solve the following global lower bound maximization of (19):

$$\max_{\mathbf{p}} \bar{\mathcal{P}}_1^{(\kappa)}(x, \mathbf{p}) \triangleq \sum_{k=1}^K \sum_{n=1}^N \frac{(x_{nk})^2}{\langle x_k^2 \rangle} \beta_{nk}^{(\kappa)}(\mathbf{p}) \tag{22a}$$

$$\begin{aligned}
 \text{s.t.} \quad &\sum_{n=1}^N x_{nk} \theta_{nk}(\mathbf{p}) \leq C_k^{\text{BH}} \langle x_k \rangle, \quad k = 1, \dots, K \\
 &(5f), (18)
 \end{aligned} \tag{22b}$$

to generate a sequence $\{p^{(\kappa)}\}$, $\kappa = 1, 2, \dots$ of feasible and improved points toward the solution of (19). At iteration κ , we use $p^{(\kappa-1)}$ as a feasible point to solve (22) and obtain $p^{(\kappa)}$. Similarly to Theorem 1, we can prove the following result.

Theorem 2: Initialized from a feasible point $p^{(0)}$, the sequence $\{p^{(\kappa)}\}$ obtained by iteratively solving (22) is a sequence of improved points, which converges to a KKT point of (19).

C. Joint Optimization of BS-UE Association and Power Allocation

The alternating optimization framework requires solving a series of convex problems (14) [cf. Section III-A] followed by solving a series of convex problems (22) [cf. Section III-B] and repeating until convergence. Realizing that solving one instant of (14) and (22) alone already provides a better point, we use them as an alternating descent to achieve a much faster convergence speed. The proposed joint optimization of BS-UE association and power allocation for sum-rate maximization is summarized in Algorithm 1. Our alternating descent approach gives flexibility in executing user association and power control in the same time slot (as in Algorithm 1) or different time slots (by selectively deactivating Step 3 or 4 of Algorithm 1).

At each iteration of Algorithm 1, the computational complexity of solving convex problems (14) and (22) is only polynomial in the number of variables and constraints. To see this, (14) can be equivalently reformulated as an optimization problem with $a \triangleq (NK + 1)$ real-valued scalar

Algorithm 1 Joint BS-UE Association and Power Allocation for Sum-Rate Maximization

- 1: Initialize $x_{nk}^{(0)} := \frac{1}{KN}$, $n = 1, \dots, N$, $k = 1, \dots, K$ and $p_k^{(0)} := P_k^{\text{max}}$, $k = 1, \dots, K$. Set $\kappa := 0$.
- 2: **repeat**
- 3: Solve convex program (14) with $p := p^{(\kappa)}$ to find optimal solution x^* .
- 4: Solve convex program (22) with $x := x^*$ to find optimal solution p^* .
- 5: Set $(x^{(\kappa+1)}, p^{(\kappa+1)}) := (x^*, p^*)$ and $\kappa := \kappa + 1$.
- 6: **until** $\left| \frac{\mathcal{P}_1(x^{(\kappa)}, p^{(\kappa)}) - \mathcal{P}_1(x^{(\kappa-1)}, p^{(\kappa-1)})}{\mathcal{P}_1(x^{(\kappa-1)}, p^{(\kappa-1)})} \right| < \epsilon$

decision variables, a linear objective, $b \triangleq (N + 2K + NK)$ linear constraints and one quadratic constraint. Similarly, (22) can be equivalently reformulated as a semidefinite program with $(2K + 1)$ scalar variables, a linear objective and a system of linear matrix inequalities. The complexity required to solve (14) and (22) is thus $\mathcal{O}((1 + a + b)a^2\sqrt{b+1})$ and $\mathcal{O}((2K + 1)^2 [(K + 1)^3 + NK^3] \sqrt{3K + NK + 1})$, respectively [37].

Theorem 3: Initialized from a feasible point $(x^{(0)}, p^{(0)})$, Algorithm 1 converges to a solution of problem (5) after a finite number of iterations for a given error tolerance $\epsilon > 0$.

Proof: The BS-UE association problem (13) and the power allocation problem (19) have the same objective function $\mathcal{P}_1(\mathbf{x}, \mathbf{p})$. From (17), (20) and (21), we have the following relations:

$$\begin{aligned}
 \mathcal{P}_1(x^{(\kappa+1)}, p^{(\kappa+1)}) &\geq \bar{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa+1)}, p^{(\kappa+1)}) \\
 &\geq \bar{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa+1)}, p^{(\kappa)}) = \mathcal{P}_1(x^{(\kappa+1)}, p^{(\kappa)}) \\
 &\geq \tilde{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa+1)}, p^{(\kappa)}) \geq \tilde{\mathcal{P}}_1^{(\kappa)}(x^{(\kappa)}, p^{(\kappa)}) \\
 &= \mathcal{P}_1(x^{(\kappa)}, p^{(\kappa)}).
 \end{aligned} \tag{23}$$

It means that alternately solving their respective convex minorants (14) and (22) always improves $\mathcal{P}_1(\mathbf{x}, \mathbf{p})$ in each iteration. As such, once initialized from a feasible point $(x^{(0)}, p^{(0)})$ that satisfies (5b), (5c) and (5f), Algorithm 1 generates a sequence $\{(x^{(\kappa)}, p^{(\kappa)})\}$ of feasible and improved points which eventually converges to a solution (\bar{x}, \bar{p}) of (5). Note that \bar{x} is a KKT point of (13) for $p = \bar{p}$ while \bar{p} is a KKT of (19) for $x = \bar{x}$. Under the stopping criterion

$$\left| \frac{\mathcal{P}_1(x^{(\kappa+1)}, p^{(\kappa+1)}) - \mathcal{P}_1(x^{(\kappa)}, p^{(\kappa)})}{\mathcal{P}_1(x^{(\kappa)}, p^{(\kappa)})} \right| < \epsilon, \tag{24}$$

Algorithm 1 terminates after a finite number of iterations for a given $\epsilon > 0$. ■

IV. PROPOSED ALTERNATING DESCENT ALGORITHM FOR MINIMUM UE RATE MAXIMIZATION

The above developed Algorithm 1 is readily extendable to solve the max-min problem (6). In this case, we consider the

following objective function.

$$\mathcal{P}_2(\mathbf{x}, \mathbf{p}) \triangleq \min_{n=1, \dots, N} \left\{ \sum_{k=1}^K \frac{\mathbf{x}_{nk}^2 r_{nk}(\mathbf{p})}{\langle \mathbf{x}_k^2 \rangle} \right\} + \lambda \sum_{k=1}^K \sum_{n=1}^N (\mathbf{x}_{nk}^2 - \mathbf{x}_{nk}), \quad (25)$$

where $\mathbf{x}_{nk} \in [0, 1]$, $n = 1, \dots, N$, $k = 1, \dots, K$ and $\lambda \geq 0$ is a constant penalty factor. The BS-UE association problem for a fixed power allocation $\mathbf{p} := p$ is now

$$\begin{aligned} \max_{\mathbf{x}_k \in [0, 1]^N} \quad & \mathcal{P}_2(\mathbf{x}, p) \\ \text{s.t.} \quad & (5b), (5c), (5e). \end{aligned} \quad (26)$$

Although the constraint set of (26) is convex, its objective is nonsmooth and nonconvex. Similar to Proposition 2, it can be shown that the following convex problem is a global lower bound maximization of (26):

$$\begin{aligned} \max_{\mathbf{x}_k \in [0, 1]^N} \quad & \tilde{\mathcal{P}}_2(\mathbf{x}, p) \triangleq \min_{n=1, \dots, N} \left\{ \sum_{k=1}^K \alpha_{nk}^{(\kappa)}(\mathbf{x}, p) \right\} \\ & + \lambda \sum_{k=1}^K \sum_{n=1}^N w_{nk} \gamma_{nk}^{(\kappa)}(\mathbf{x}), \\ \text{s.t.} \quad & (5b), (5c), (5e), \end{aligned} \quad (27)$$

where $\alpha_{nk}^{(\kappa)}(\mathbf{x}, p)$ and $\gamma_{nk}^{(\kappa)}(\mathbf{x})$ have previously been defined in (15).

Next, the power allocation problem for a fixed $\mathbf{x} := x$ is

$$\begin{aligned} \max_{\mathbf{p}} \quad & \mathcal{P}_2(x, \mathbf{p}) \triangleq \min_{k=1, \dots, K} \left\{ \sum_{n=1}^N \frac{x_{nk}^2}{\langle x_k^2 \rangle} r_{nk}(\mathbf{p}) \right\} \\ \text{s.t.} \quad & (5e), (5f), \end{aligned} \quad (28)$$

which has a nonsmooth nonconvex objective function and a nonconvex set. Similar to Proposition 3, it can be shown that the following convex problem is a global lower bound maximization of (28):

$$\begin{aligned} \max_{\mathbf{p}} \quad & \bar{\mathcal{P}}_2^{(\kappa)}(x, \mathbf{p}) \triangleq \min_{k=1, \dots, K} \left\{ \sum_{n=1}^N \frac{x_{nk}^2}{\langle x_k^2 \rangle} \beta_{nk}(\mathbf{p}) \right\} \\ \text{s.t.} \quad & (22b), (5f), \end{aligned} \quad (29)$$

where $\beta_{nk}(\mathbf{p})$ has previously been defined in (21).

To solve problem (6) in *both* \mathbf{x} and \mathbf{p} , we modify Algorithm 1 as follows. In Step 3, we solve convex problem (27) instead of (14). In Step 4, we solve convex problem (29) instead of (22). And because the objective function is now $\mathcal{P}_2(x, \mathbf{p})$, the proposed algorithm for problem (6) terminates when $\left| \frac{\mathcal{P}_2(x^{(\kappa)}, \mathbf{p}^{(\kappa)}) - \mathcal{P}_2(x^{(\kappa-1)}, \mathbf{p}^{(\kappa-1)})}{\mathcal{P}_2(x^{(\kappa-1)}, \mathbf{p}^{(\kappa-1)})} \right| < \epsilon$. We shall refer to this modified algorithm as Algorithm 2.

In each iteration of Algorithm 2, the computational complexity of solving problem (27) is $\mathcal{O}((N + a + d)c^2\sqrt{c + N})$, because (27) can be equivalently reformulated as an optimization problem with $a = (NK + 1)$ scalar real decision variables, a linear objective, $c \triangleq (N + K)$ linear constraints and N quadratic constraints [37]. Similarly, the complexity of solving problem (29) is $\mathcal{O}([cd^2(K + 1)^2 + d^2NK^3]\sqrt{3K + NK + N})$, because (29) can be reformulated as a semidefinite program with $d \triangleq (2K + 1)$ scalar variables, a linear objective and a system

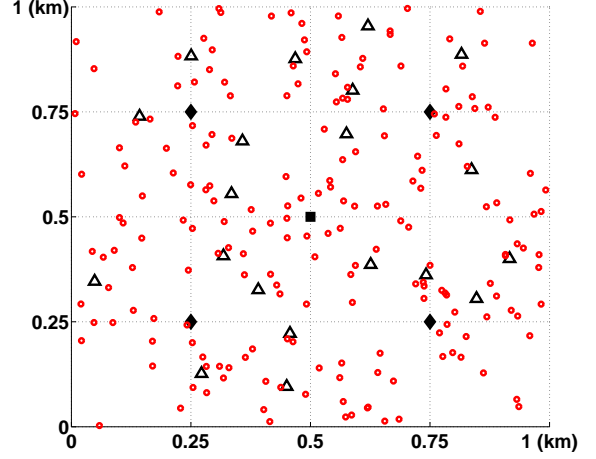


Fig. 2. A three-tier network with one fixed MBS (black square), four fixed PBSs (black diamonds), twenty random FBSs (black triangles) and 200 random UEs (red circles)

of linear matrix inequalities. Finally, similar to Theorem 3, it can be proved that once initialized from a feasible point $(x^{(0)}, p^{(0)})$ that satisfies (5b), (5c) and (5f), Algorithm 2 converges after a finite number of iterations for a given error tolerance ϵ .

V. ILLUSTRATIVE EXAMPLES

Consider a three-tier HetNet where four pico BSs (PBSs) and twenty femto BSs (FBSs) are deployed within a macrocell of size $1,000\text{m} \times 1,000\text{m}$. The locations of MBS and PBSs are fixed whereas those of FBSs are random, as shown in Fig. 2. We assume there are $N = 200$ UEs randomly distributed over the macrocell coverage area. The network topology is then fixed during the optimization process. Without loss of generality, we only consider the effect of pathloss when generating the channel gains. To illustrate the impact of imperfect backhaul links, we assume that the MBS, PBSs, and FBSs are each equipped with a backhaul link of capacity C^{BH} , $C^{\text{BH}}/3$ and $C^{\text{BH}}/10$, respectively. Following [38], we choose $C^{\text{BH}} \in \{100, 150, 200, \infty\}$ Mbps where $C^{\text{BH}} = \infty$ represents the ideal backhaul. For simplicity, we set the required minimum UE throughput as $R_n^{\text{QoS}} = R^{\text{QoS}}$, $n = 1, \dots, N$. The error tolerance for the algorithms is set as $\epsilon = 10^{-4}$. Other 3GPP LTE parameters used to setup our simulations are listed in Table I [39]. Note that we divide the obtained rate results by $\ln(2)$ to arrive at the unit of bps/Hz.

First, we compare the sum-rate performance of the joint design in Algorithm 1 to that of Algorithm 1 but with full BS transmit power (i.e., no power control). We use the heuristic BS-UE association schemes, namely, max-SINR and DCD [10] as benchmarks where full BS transmit power is also assumed. As both benchmark schemes assume ideal backhaul, we set $C^{\text{BH}} = \infty$ here for a fair comparison. And since the max-SINR and DCD schemes do not include the minimum UE throughput constraint, we first assume $R^{\text{QoS}} = 0$ in these two schemes to find their BS-UE associations, followed by

TABLE I
SIMULATION PARAMETERS USED IN ALL NUMERICAL EXAMPLES

Parameter	Value
Minimum distance between MBS-UE	35m
Minimum distance between PBS/FBS-UE	10m
Path loss model for MBS-UE links	$128.1 + 37.6 \ln_{10}(d)$, d is in km
Path loss model for PBS-UE links and FBS-UE links	$140.7 + 36.7 \ln_{10}(d)$, d is in km
Maximum MBS transmit power	43dBm
Maximum PBS transmit power	24dBm
Maximum FBS transmit power	20dBm
Background noise power	-104dBm
System bandwidth	10MHz
Frequency reuse factor	1

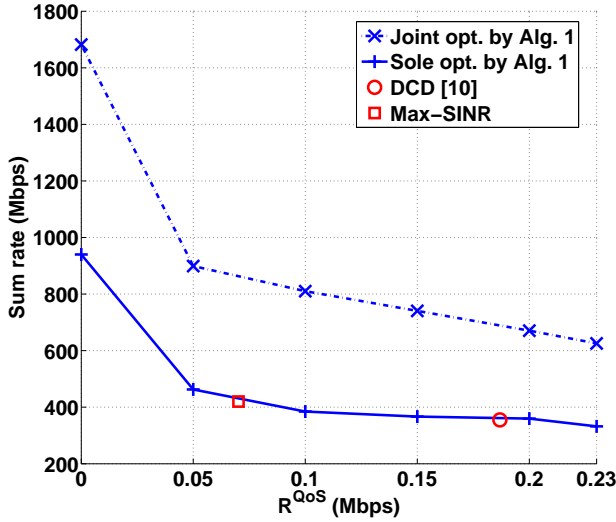


Fig. 3. Sum-rate performance of Algorithm 1 under ideal backhaul links.

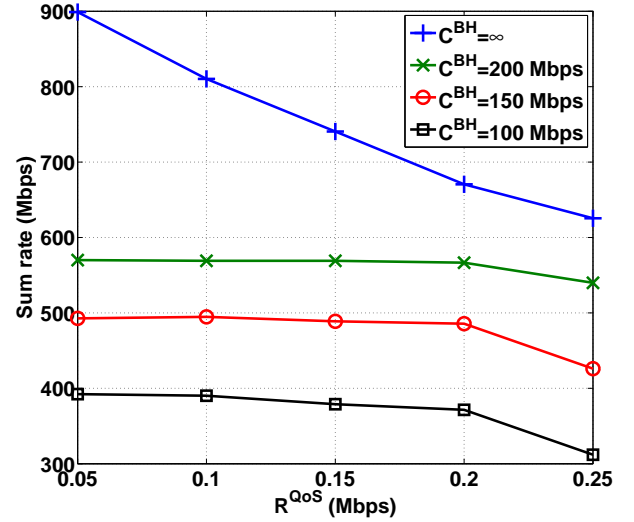


Fig. 4. Effects of QoS constraints and limited backhaul capacity on the sum-rate performance of Algorithm 1

calculating their achieved sum-rates and minimum UE rates. Fig. 3 shows that joint design of load balancing and interference management in Algorithm 1 gives much higher network throughput over load balancing alone. This can be explained by noting that the joint design has an extra dimension of BS transmit power to optimize to further enhance the sum-rate performance. When comparing Algorithm 1 with full BS transmit power to the max-SINR and DCD schemes, the former offers more flexibility in setting the desired minimum R^{QoS} . Furthermore, for the same values of R^{QoS} that are achieved by the max-SINR and DCD schemes, the sole BS-UE design by Algorithm 1 gives a slightly better sum-rate performance as can be observed from Fig. 3.

Next, we evaluate the effects of QoS constraints and limited backhaul capacity in the joint design of Algorithm 1. Fig. 4 shows that as we move away from the assumption of ideal backhaul, the total throughput is gradually degraded. This observation is as expected because the feasible region of problem (5) becomes more restricted. For each value of C^{BH} , Fig. 4 also shows that lowering the UE throughput requirement R^{QoS} actually increases the total throughput. However, while

such an throughput improvement is pronounced for the ideal backhaul, it is not much so for limited backhaul capacity cases where reducing R^{QoS} beyond 0.2Mbps only marginally improves the sum-rate. Our numerical analysis reveals that most of the PBSs and FBSs have fully utilized their respective limited backhaul capacities of $C^{\text{BH}}/3$ and $C^{\text{BH}}/10$, leaving no room for further throughput improvement at these small cells even if R^{QoS} is small. Such an observation again verifies that backhaul capacity is in fact a bottleneck for network performance.

Fig. 5 further demonstrates that for a given R^{QoS} , switching from ideal backhaul ($C^{\text{BH}} = \infty$) to non-ideal backhaul ($C^{\text{BH}} = 100\text{Mbps}$) may limit the offloading capability of small cells. Indeed, many UEs will be transferred from the small cells back to the macrocell. This observation can be explained as follows. The small cells are more easily overloaded in the non-ideal backhaul case as their backhaul capacity is much smaller than that of the macrocell. To meet their backhaul limitations, small-cell BSs decrease their transmit power to shrink their cell size and serve fewer UEs with lower cell throughput.

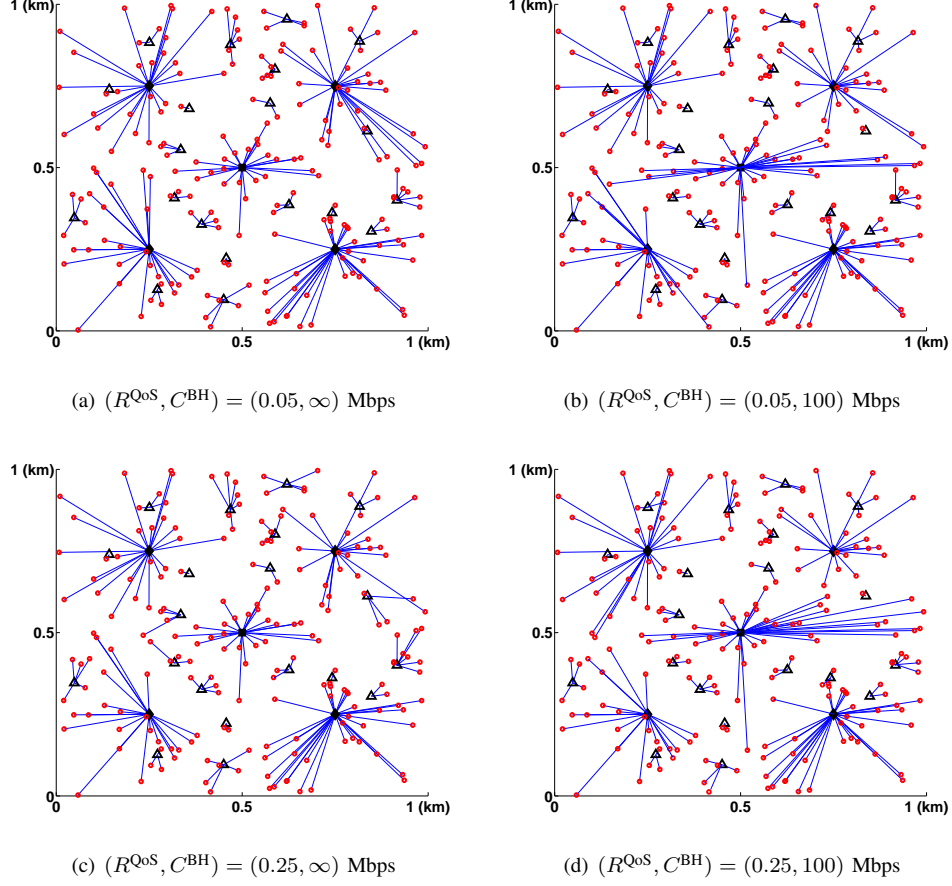


Fig. 6. Changes in BS-UE associations by Algorithm 1 under various choices of UE QoS requirements and backhaul capacity

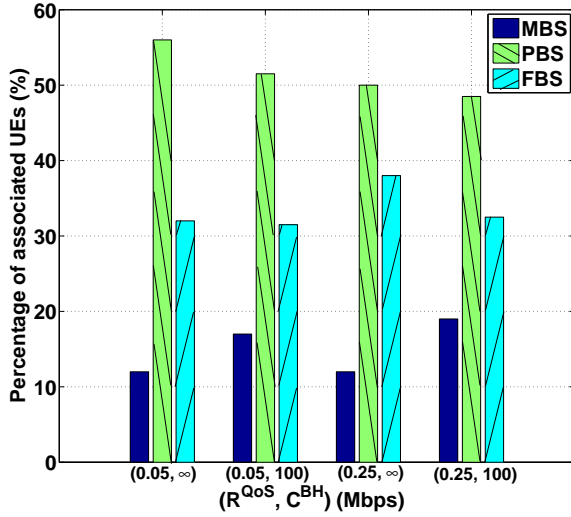


Fig. 5. Effects of QoS constraints and limited backhaul capacity on the load distribution by Algorithm 1

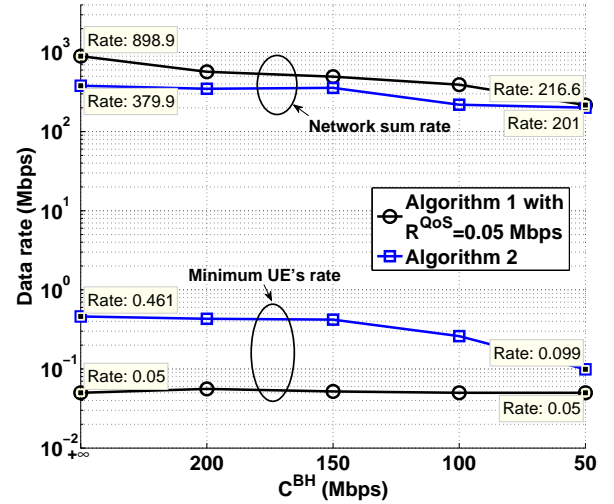


Fig. 7. Fairness by Algorithm 2 and effects of limited backhaul capacity on the minimum UE rate

With 19dB-23dB higher in power budget compared to that available to small-cell BSs, the MBS then increases its transmit power (and effectively its coverage area) to take over the UEs pushed out by the small cells. And with 3-10 times more

backhaul capacity, the MBS is still able to accommodate the incoming traffic. This can be best observed in Fig. 6(b) [cf. Fig. 6(a)] and Fig. 6(d) [cf. Fig. 6(c)], where the MBS serves more distant UEs for $C^{\text{BH}} = 100$ Mbps.

Fig. 7 demonstrates the fairness given by Algorithm 2 for

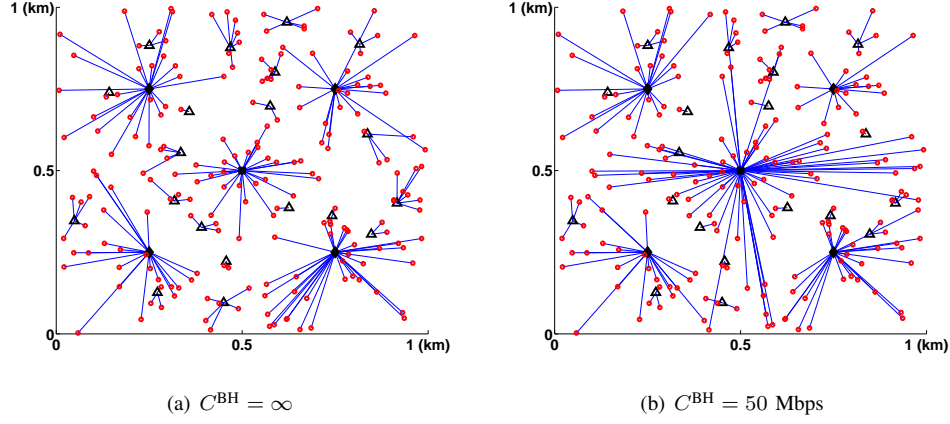


Fig. 8. More UEs switch to MBS when backhaul capacity is limited in Algorithm 2

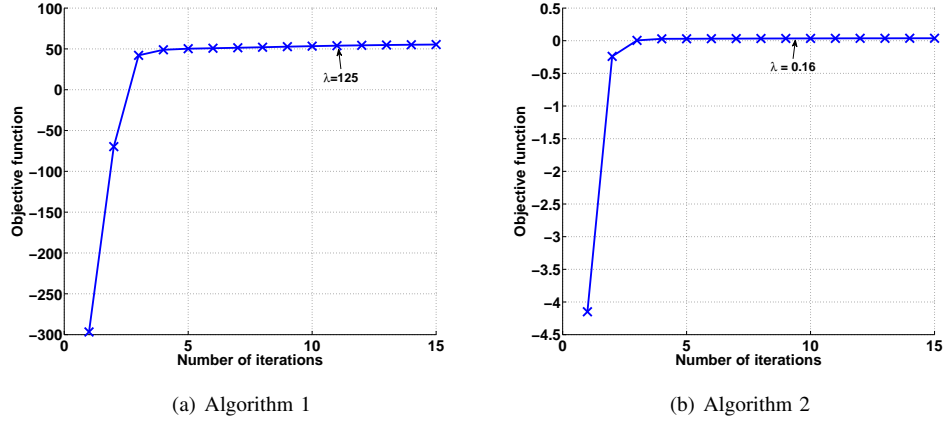


Fig. 9. Convergence of proposed algorithms for $\epsilon = 10^{-4}$

max-min UE rates. Here, we compare against Algorithm 1 for sum-rate maximization where $R^{\text{QoS}} = 0.05 \text{ Mbps}$ is assumed. As seen from the figure, Algorithm 2 improves the minimum UE rate as the cost of reduced total throughput. Furthermore, for $C^{\text{BH}} \geq 150 \text{ Mbps}$, reducing backhaul capacity does not affect much the minimum UE rate and sum-rate performance of Algorithm 2 because the ample backhaul capacity at each cell can still accommodate more data traffic. However, for $C^{\text{BH}} \leq 100 \text{ Mbps}$, the minimum UE rate by Algorithm 2 starts to fall dramatically. And at $C^{\text{BH}} = 50 \text{ Mbps}$, the achieved minimum UE rate drops by more than 4.5 times compared to that in the ideal backhaul case. At this point, a significantly larger number of UEs turns to the MBS for service as shown Fig. 8. This is because the MBS still has available backhaul capacity while the PBSs and FBSs are more likely to be overloaded.

Finally, we examine the convergence of the proposed algorithms. It is sufficient to choose a large value of the penalty factor λ in the exact penalty method. To improve the convergence speed of Algorithms 1 and 2, our implementation starts with $\lambda = 10^3$ and fine-tunes λ through a bisection search until the objective functions no longer change and binary values of \mathbf{x} are found. For brevity, only the case of $C^{\text{BH}} = 100 \text{ Mbps}$ is presented for illustration. Fig. 9(a) plots the convergence of the

objective function (19) by Algorithm 1 for $R^{\text{QoS}} = 0.1 \text{ Mbps}$. Fig. 9(b) plots the convergence of the objective function (25) by Algorithm 2. In these plots, the system bandwidth is normalized to unit to ensure the compatibility of the utility function and the penalty term in (19) and (25). The number of iterations in each plot corresponds to the presented values of λ . As can be seen from Fig. 9, the proposed algorithms only require at most ten iterations to converge. It is worth noting that each iteration of our algorithms involves solving only two easy convex problems, each with polynomial complexity.

VI. CONCLUSIONS

In this paper, we have proposed new joint BS-UE association and power control schemes for HetNets. Specifically, we have addressed two difficult mixed-integer optimization problems: (i) sum-throughput maximization under QoS constraints and (ii) maximization of minimum UE throughput. Our problem formulations also include the practical constraint of limited backhaul capacity. Our developed alternative descent algorithms are based on an exact penalty method combined with successive convex programming, where we address the binary BS-UE association problem and the nonconvex power allocation problem separately. At each iteration, only two simple convex problems are solved in the same time scale. Our

algorithms improve the objective functions in each iteration and converge eventually. Simulation results have demonstrated the usefulness of our devised algorithms in both traffic offload-ing and interference management.

APPENDIX A PROOF OF PROPOSITION 2

Firstly, notice that functions

$$f_{nk}(\mathbf{x}_{nk}, \mathbf{y}_k) \triangleq \frac{r_{nk}(p)\mathbf{x}_{nk}^2}{\mathbf{y}_k}, \quad k = 1, \dots, K$$

are jointly convex in \mathbf{x}_{nk} and \mathbf{y}_k . Therefore, $f_{nk}(\mathbf{x}_{nk}, \mathbf{y}_k)$ admits its first order approximation at $(x_{nk}^{(\kappa)}, y_k^{(\kappa)})$ as a lower bound [40] as follows

$$\begin{aligned} f_{nk}(\mathbf{x}_{nk}, \mathbf{y}_k) &\geq f_{nk}(x_{nk}^{(\kappa)}, y_k^{(\kappa)}) + \frac{2r_{nk}(p)x_{nk}^{(\kappa)}(\mathbf{x}_{nk} - x_{nk}^{(\kappa)})}{y_k^{(\kappa)}} \\ &\quad - \frac{r_{nk}(p)(x_{nk}^{(\kappa)})^2}{(y_k^{(\kappa)})^2}(\mathbf{y}_k - y_k^{(\kappa)}). \end{aligned}$$

By replacing $\mathbf{y}_k := \langle \mathbf{x}_k^2 \rangle > 0$, $\forall k = 1, \dots, K$, we have

$$f_{nk}(\mathbf{x}_{nk}, [\langle \mathbf{x}_k^2 \rangle]_{k=1, \dots, K}) = \frac{r_{nk}(p)\mathbf{x}_{nk}^2}{\langle \mathbf{x}_k^2 \rangle} \geq \alpha_{nk}^{(\kappa)}(\mathbf{x}, p).$$

Secondly, since $\mathbf{x}_{nk}^2 - \mathbf{x}_{nk}$ is a convex quadratic function, it also admits its first order approximation at $x^{(\kappa)}$ as a lower bound as

$$\begin{aligned} \mathbf{x}_{nk}^2 - \mathbf{x}_{nk} &\geq (x_{nk}^{(\kappa)})^2 - x_{nk}^{(\kappa)} + (2x_{nk}^{(\kappa)} - 1)(\mathbf{x}_{nk} - x_{nk}^{(\kappa)}) \\ &= \gamma_{nk}^{(\kappa)}(\mathbf{x}). \end{aligned}$$

Therefore, we have

$$\mathcal{P}_1(\mathbf{x}, p) \geq \tilde{\mathcal{P}}_1^{(\kappa)}(\mathbf{x}, p)$$

at a point $(x^{(\kappa)}, p)$. This completes the proof.

APPENDIX B PROOF OF PROPOSITION 3

To prove Proposition 3, we need some preliminary results first. Let us define

$$g_n(x_1, \dots, x_n) \triangleq \left(\sum_{i=1}^n x_i^{-1} \right)^{-1}. \quad (30)$$

Theorem 4: For $x_1 > 0$, $x_2 > 0$, $g_2(x_1, x_2)$ is a concave and monotonically increasing function in (x_1, x_2) .

Proof: One has

$$g_2(x_1, x_2) = x_1 - x_1^2/(x_1 + x_2)$$

which is concave in $x_1 > 0$, $x_2 > 0$. Moreover, as $x_1^{-1} + x_2^{-1}$ is monotonically decreasing in $x_1 > 0$, $x_2 > 0$, $g_2(x_1, x_2) = 1/(x_1^{-1} + x_2^{-1})$ is monotonically increasing in $x_1 > 0$, $x_2 > 0$. ■

Theorem 5: The function $g_n(x_1, \dots, x_n) = (\sum_{i=1}^n x_i^{-1})^{-1}$ is concave in $x_i > 0$, $i = 1, \dots, n$, $\forall n \geq 1$.

Proof: In Theorem 4, we have proved that $g_2(x_1, x_2)$ is concave in $x_1 > 0$, $x_2 > 0$. Assuming that

$g_{n-1}(x_1, \dots, x_{n-1})$ is concave in $x_i > 0$, $i = 1, \dots, n-1$ for some n , we now show that $g_n(x_1, \dots, x_n)$ is also concave. Firstly, one has

$$\begin{aligned} g_n(x_1, \dots, x_n) &= \left(\sum_{i=1}^n x_i^{-1} \right)^{-1} \\ &= x_n - x_n^2(x_n + g_{n-1}(x_1, \dots, x_{n-1}))^{-1} \\ &= g_2(x_n, g_{n-1}(x_1, \dots, x_{n-1})). \end{aligned}$$

Since $g_2(x, y)$ is a concave and monotonically increasing function in (x, y) and $g_{n-1}(x_1, \dots, x_{n-1})$ is assumed to be concave, we have

$$\begin{aligned} g_n(\alpha(x_1, \dots, x_n) + \beta(y_1, \dots, y_n)) &= g_2(\alpha x_n + \beta y_n, g_{n-1}(\alpha(x_1, \dots, x_{n-1}) \\ &\quad + \beta(y_1, \dots, y_{n-1}))) \\ &\geq g_2(\alpha x_n + \beta y_n, \alpha g_{n-1}(x_1, \dots, x_{n-1}) \\ &\quad + \beta g_{n-1}(y_1, \dots, y_{n-1})) \\ &\geq \alpha g_2(x_n, g_{n-1}(x_1, \dots, x_{n-1})) \\ &\quad + \beta g_2(y_n, g_{n-1}(y_1, \dots, y_{n-1})) \\ &= \alpha g_n(x_1, \dots, x_n) + \beta g_n(y_1, \dots, y_n) \end{aligned}$$

for $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$ and $\forall x_i > 0, y_i > 0, i = 1, \dots, n > 0$. Thus, $g_n(x_1, \dots, x_n)$ is also concave in $x_i > 0$, $i = 1, \dots, n$, $\forall n \geq 2$. ■

Theorem 6: The function $\ln(\sum_{i=1}^n x_i^{-1})$ is convex in $x_i > 0$, $i = 1, \dots, n$, $\forall n \geq 1$.

Proof: Let us define $f(x) \triangleq \ln(x)$. Then $\ln(\sum_{i=1}^n x_i^{-1}) = -f(g_n(x_1, \dots, x_n))$. Since $g_n(x_1, \dots, x_n)$ is a concave function in $x_i > 0, i = 1, \dots, n$, according to Theorem 5 we have

$$\begin{aligned} g_n(\alpha(x_1, \dots, x_n) + \beta(y_1, \dots, y_n)) &\geq \alpha g_n(x_1, \dots, x_n) + \beta g_n(y_1, \dots, y_n), \end{aligned}$$

with $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$ and $\forall x_i > 0, y_i > 0, i = 1, \dots, n > 0$. Moreover, since $f(x)$ is a concave and monotonically increasing function, we have

$$\begin{aligned} f(g_n(\alpha(x_1, \dots, x_n) + \beta(y_1, \dots, y_n))) &\geq f(\alpha g_n(x_1, \dots, x_n) + \beta g_n(y_1, \dots, y_n)) \\ &\geq \alpha f(g_n(x_1, \dots, x_n)) + \beta f(g_n(y_1, \dots, y_n)). \end{aligned}$$

This shows the concavity of $f(g_n(x_1, \dots, x_n))$. Therefore, $-f(g_n(x_1, \dots, x_n))$ is a convex function. ■

We are now ready to prove Proposition 3. One can write

$$r_{nk}(\mathbf{p}) \triangleq \ln \left(\sum_{j=1}^K g_{nj} \mathbf{p}_j + \sigma^2 \right) - \ln \left(\sum_{j \neq k, j=1}^K g_{nj} \mathbf{p}_j + \sigma^2 \right).$$

Since $\ln(x)$ is a concave function with $x > 0$, it is true that

$$\begin{aligned} \ln \left(\sum_{j=1}^K g_{nj} \mathbf{p}_j + \sigma^2 \right) &\leq \ln \left(\sum_{j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2 \right) \\ &\quad + \frac{\sum_{j=1}^K g_{nj}(\mathbf{p}_j - p_j^{(\kappa)})}{\sum_{j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2}, \quad (31) \end{aligned}$$

at some point $p^{(\kappa)}$.

In addition, upon defining $f(x) \triangleq \ln(\sum_{i=1}^n x_i^{-1})$ with $x = [x_1, \dots, x_n]$, $x_i > 0$, $\forall i$, one has

$$f(x) \geq \ln\left(\sum_{i=1}^n (x_i^{(\kappa)})^{-1}\right) - \frac{1}{\sum_{i=1}^n (x_i^{(\kappa)})^{-1}} \sum_{i=1}^n \left(\frac{x_i - x_i^{(\kappa)}}{(x_i^{(\kappa)})^2}\right)$$

at some $x^{(\kappa)}$, due to the convexity of $f(x)$ by Theorem 6. Thus, one also has the following bound

$$\ln\left(\sum_{i=1}^n x_i\right) \geq \ln\left(\sum_{i=1}^n x_i^{(\kappa)}\right) - \frac{1}{\sum_{i=1}^n x_i^{(\kappa)}} \sum_{i=1}^n (x_i^{(\kappa)})^2 \left(\frac{1}{x_i} - \frac{1}{x_i^{(\kappa)}}\right).$$

It follows that

$$\begin{aligned} & \ln\left(\sum_{j \neq k, j=1}^K g_{nj} \mathbf{p}_j + \sigma^2\right) \\ & \leq \ln\left(\sum_{j \neq k, j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2\right) - \frac{1}{\sum_{j \neq k, j=1}^K g_{nj} p_j^{(\kappa)} + \sigma^2} \\ & \quad \times \sum_{j \neq k, j=1}^K \left(g_{nj} p_j^{(\kappa)}\right)^2 \left(\frac{1}{g_{nj} \mathbf{p}_j} - \frac{1}{g_{nj} p_j^{(\kappa)}}\right). \quad (32) \end{aligned}$$

By combining (31) and (32), (20) follows. Similarly, (21) can be proved with minor modifications to (31) and (32).

REFERENCES

- [1] J. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.
- [2] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun. Mag.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.
- [3] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.
- [4] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K -tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.
- [5] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.
- [6] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.
- [7] S. Lee, K. Son, H. Gong, and Y. Yi, "Base station association in wireless cellular networks: An emulation based approach," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2720–2729, Aug. 2012.
- [8] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.
- [9] D. Fouladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 248–257, Jan. 2013.
- [10] K. Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1100–1113, Jun. 2014.
- [11] S. H. Low and D. E. Lapsley, "Optimization flow control. I. Basic algorithm and convergence," *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, Dec. 1999.
- [12] H. Boostanimehr and V. Bhargava, "Unified and distributed QoS-driven cell association algorithms in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1650–1662, Mar. 2015.
- [13] S. Corroy and R. Mathar, "Semidefinite relaxation and randomization for dynamic cell association in heterogeneous networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 2373–2378.
- [14] S. Corroy, L. Falconetti, and R. Mathar, "Cell association in small heterogeneous networks: Downlink sum rate and min rate maximization," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, Apr. 2012, pp. 888–892.
- [15] D. Liu, Y. Wang, Y. Chen, M. ElKashlan, K. K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *IEEE Communications Surveys and Tuts.*, vol. 18, no. 2, pp. 1018–1044, Second quarter 2016.
- [16] J. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in HetNets: old myths and open problems," *IEEE Wireless Commun. Mag.*, vol. 21, no. 2, pp. 18–25, Apr. 2014.
- [17] H. Beyranvand, W. Lim, M. Maier, C. Verikoukis, and J. Salehi, "Backhaul-aware user association in FiWi enhanced LTE-A heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 2992–3003, Jun. 2015.
- [18] J. Ghimire and C. Rosenberg, "Revisiting scheduling in heterogeneous networks when the backhaul is limited," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2039–2051, Oct. 2015.
- [19] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014.
- [20] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.
- [21] J. Yu, Y. Wang, X. Lin, and Q. Zhang, "Power allocation for CoMP system with backhaul limitation," in *Proc. IEEE Intl. Conf. Commun. (ICC)*, Jun. 2014, pp. 4759–4764.
- [22] V. N. Ha and L. B. Le, "Distributed base station association and power control for heterogeneous cellular networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 282–296, Jan. 2014.
- [23] R. Sun, M. Hong, and Z.-Q. Luo, "Joint downlink base station association and power control for max-min fairness: Computation and complexity," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1040–1054, Jun. 2015.
- [24] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1479–1489, Dec. 2010.
- [25] M. Hong and A. Garcia, "Mechanism design for base station association and resource allocation in downlink OFDMA network," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2238–2250, Dec. 2012.
- [26] D. P. Bertsekas, *Nonlinear Programming*, 2nd. Athena Scientific, 1999.
- [27] A. H. Jafari, D. López-Pérez, H. Song, H. Clausen, L. Ho, and J. Zhang, "Small cell backhaul: challenges and prospective solutions," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–18, 2015. [Online]. Available: <http://dx.doi.org/10.1186/s13638-015-0426-y>
- [28] G. Zhang, T. Q. S. Quek, M. Kountouris, A. Huang, and H. Shan, "Fundamentals of heterogeneous backhaul design - analysis and optimization," *IEEE Trans. Commun.*, no. 2, pp. 876–889, Feb. 2016.
- [29] J. Olmos, R. Ferrus, and H. Galeana-Zapien, "Analytical modeling and performance evaluation of cell selection algorithms for mobile networks with backhaul capacity constraints," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6011–6023, Dec. 2013.
- [30] Qian, C. Yang, and A. Molisch, "Downlink base station cooperative transmission under limited-capacity backhaul," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3746–3759, Aug. 2013.
- [31] Q. Kuang, J. Speidel, and H. Droste, "Joint base-station association, channel assignment, beamforming and power control in heterogeneous networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5.
- [32] K. Chitti, Q. Kuang, and J. Speidel, "Joint base station association and power allocation for uplink sum-rate maximization," in *Proc. IEEE Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Jun. 2013, pp. 6–10.
- [33] E. Che, H. D. Tuan, and H. Nguyen, "Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5481–5495, Oct. 2014.

- [34] J. F. Bonnans, J. C. Gilbert, C. Lemarechal, and C. Sagastizabal, *Numerical Optimization Theoretical and Practical Aspects*, 2nd ed. Springer, 2006.
- [35] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, Jul. 1978.
- [36] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local d.c. programming," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 510–515, Feb. 2012.
- [37] A. Nemirovski, "Interior point polynomial time methods in convex programming," 2004. [Online]. Available: http://www2.isye.gatech.edu/~nemirovs/Lect_IPM.pdf
- [38] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan, "The evolution of cellular backhaul technologies: Current issues and future trends," *IEEE Commun. Surveys and Tutorials*, vol. 13, no. 1, pp. 97–113, First Quarter 2011.
- [39] 3GPP TS 36.814 V9.0.0, "3GPP technical specification group radio access network, evolved universal terrestrial radio access (E-UTRA): Further advancements for E-UTRA physical layer aspects (Release 9)," 2010.
- [40] H. Tuy, *Convex Analysis and Global Optimization*. Kluwer Academic, 1998.



Ho Huu Minh Tam was born in Ho Chi Minh City, Vietnam. He received the B.S. degree in Electrical Engineering and Telecommunications from Ho Chi Minh City University of Technology in 2012. He is currently working toward the Ph.D degree at the University of Technology Sydney, Sydney, N.S.W., Australia, under the supervision of Prof. Tuan Hoang. His research interest is in optimization techniques in signal processing for wireless communications.



Hoang Duong Tuan received the Diploma (Hons.) and Ph.D. degrees in applied mathematics from Odessa State University, Ukraine, in 1987 and 1991, respectively. He spent nine academic years in Japan as an Assistant Professor in the Department of Electronic-Mechanical Engineering, Nagoya University, from 1994 to 1999, and then as an Associate Professor in the Department of Electrical and Computer Engineering, Toyota Technological Institute, Nagoya, from 1999 to 2003. He was a Professor with the School of Electrical Engineering and Telecommunications, University of New South Wales, from 2003 to 2011. He is currently a Professor with the Centre for Health Technologies, University of Technology Sydney. He has been involved in research with the areas of optimization, control, signal processing, wireless communication, and biomedical engineering for more than 20 years.



Duy Trong Ngo (S'08-M'15) received the B.Eng. (with First-class Honours and the University Medal) degree in telecommunication engineering from The University of New South Wales, Australia in 2007, the M.Sc. degree in electrical engineering (communication) from University of Alberta, Canada in 2009, and the Ph.D. degree in electrical engineering from McGill University, Canada in 2013.

Since 2013, he has been a Senior Lecturer with the School of Electrical Engineering and Computing, The University of Newcastle, Australia, where he currently leads the research effort in design and optimization for 5G wireless communications networks.

In 2013, Dr. Ngo was awarded two prestigious Postdoctoral Fellowships from the Natural Sciences and Engineering Research Council of Canada and the Fonds de recherche du Québec – Nature et technologies. At The University of Newcastle, he received the 2015 Vice-Chancellor's Award for Research and Innovation Excellence and the 2015 Pro Vice-Chancellor's Award for Research Excellence in the Faculty of Engineering and Built Environment.



Trung Q. Duong (S'05, M'12, SM'13) received his Ph.D. degree in Telecommunications Systems from Blekinge Institute of Technology (BTH), Sweden in 2012. Since 2013, he has joined Queen's University Belfast, UK as a Lecturer (Assistant Professor). His current research interests include physical layer security, energy-harvesting communications, cognitive relay networks. He is the author or co-author of more than 220 technical papers published in scientific journals (110 articles) and presented at international conferences (110 papers).

Dr. Duong currently serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, IET COMMUNICATIONS. He was a Editor of WILEY TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES, ELECTRONICS LETTERS and has also served as the Guest Editor of the special issue on some major journals including IEEE JOURNAL IN SELECTED AREAS ON COMMUNICATIONS, IET COMMUNICATIONS, IEEE ACCESS, IEEE WIRELESS COMMUNICATIONS MAGAZINE, IEEE COMMUNICATIONS MAGAZINE, EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING, EURASIP JOURNAL ON ADVANCES SIGNAL PROCESSING. He was awarded the Best Paper Award at the IEEE Vehicular Technology Conference (VTC-Spring) in 2013, IEEE International Conference on Communications (ICC) 2014. He is the recipient of prestigious Royal Academy of Engineering Research Fellowship (2016-2021).



Vincent Poor (S72, M77, SM82, F87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering. From 2006 till 2016, he served as Dean of Princetons School of Engineering and Applied Science. Dr. Poors research interests are in the areas of statistical signal processing, stochastic analysis and information theory,

and their applications in wireless networks and related fields. Among his publications in these areas is the recent book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and a foreign member of the Royal Society. He is also a Fellow of the American Academy of Arts and Sciences and the National Academy of Inventors, and of other national and international academies. He received the Technical Achievement and Society Awards of the IEEE Signal Processing Society in 2007 and 2011, respectively. Recent recognition of his work includes the 2014 URSI Booker Gold Medal, the 2015 EURASIP Athanasios Papoulis Award, the 2016 John Fritz Medal, and honorary doctorates from Aalborg University, Aalto University, HKUST and the University of Edinburgh.